Automatic reconstruction of wall features under clutter and occlusion

Michailidis, Georgios-Tsampikos ; Pajarola, Renato

Abstract: In this paper, a new method to extract wall openings (windows and doors) in interior scenes from point clouds under cluttered and occluded environments is presented. For each wall surface or a room represented by a bounding polyhedron and its 3D scan points, our method constructs a planar cell complex representation, which is used for the wall features segmentation using a graph-cut method. We evaluate our results of the proposed approach on real-world 3D scans of indoor environments and demonstrate its validity.

Noname manuscript No.
(will be inserted by the editor)

Automatic Reconstruction of Wall Features under Clutter and Occlusion

Georgios-Tsampikos Michailidis · Renato Pajarola

Received: date / Accepted: date

Abstract In this paper, a new method to extract wall openings (windows and doors) in interior scenes from point clouds under cluttered and occluded environments is presented. For each wall surface or a room represented by a bounding polyhedron and its 3D scan points, our method constructs a planar cell complex representation, which is used for the wall features segmentation using a graph-cut method. We evaluate our results of the proposed approach on real-world 3D scans of indoor environments and demonstrate its validity.

Keywords scene reconstruction · point cloud processing · LiDAR reconstruction · wall surface reconstruction · wall openings detection

1 Introduction

In recent years, an increased demand for semantically rich 3D information from building models emerged. Interior building reconstruction is an active research field, but yet less advanced than the better studied case of reconstructing the exterior outlines and façades of buildings [1,2], mainly because of the presence of objects and obstacles, or the increased clutter and occlusions, which prevent a clean data acquisition and reconstruction.

A number of methods have recently been proposed to overcome these limitations and reconstruct the structural elements of indoor scenes [1–12]. Despite the significant improvements towards automation in architectural room elements reconstruction [1,3,9], some techniques are based on specialized equipment [4,6], which limits their application field and usage, or depend on the Manhattan-world assumption which is often too restrictive [5,7]. Furthermore, fully automated 3D room modeling from interior scans has been achieved only under ideal conditions and lack of occlusions [7,13], and the capturing and incorporation of reliable as-is data remains still very challenging considering the strict constraints in accuracy, time and cost [4,8].

For identification of wall openings (e.g. windows and doors) in indoor environments, several methods have been developed [1,3,14–16], either relying on analyzing data density [14], or by using machine learning and probabilistic techniques [15,16]. Spatial and functional relationships were also explored [1,3], modeling the main structural components of room interiors despite the presence of clutter and occlusion. Recent studies try also to combine data from laser scanners and cameras, in order to avoid the challenging window detections from point clouds and the above mentioned restrictions [2,15–17].

In this paper we propose a new method capable of automatically recovering the architectural components of wall features in indoor environments, i.e. windows and doors, under the presence of significant amounts of clutter and occlusion, without requiring any tuning or additional imagery or depth data, and avoiding the restrictions of the aforementioned techniques.

We also demonstrate the validity of our approach by presenting results from both window and door reconstructions on real-world 3D scans of typical indoor environments.

2 Feature Outlines

We formulate the wall features extraction as a graph-cut optimization problem on a 2D cell complex defined by the line features on the reconstructed wall surface. A

G.-T. Michailidis, R. Pajarola
Dept. of Informatics, University of Zurich, Switzerland
E-mail: {gtmichail,pajarola}@ifi.uzh.ch

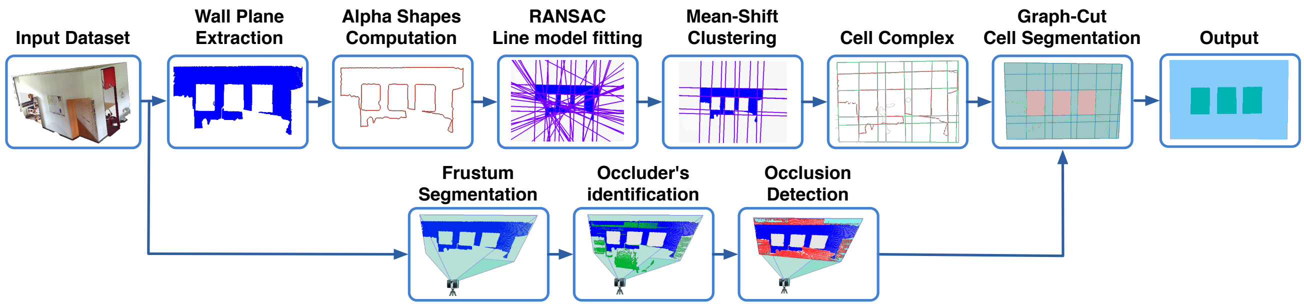


Fig. 1: Block diagram of the proposed reconstruction technique.

general diagram presenting the pipeline process of our proposed method is shown in Fig. 1.

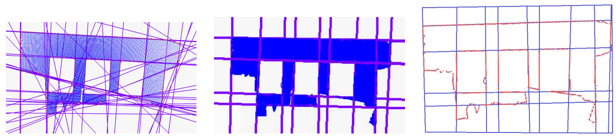
First, the scene of interest is scanned and the acquired input scans are registered and merged together to output a semi-structured 3D point cloud. Then, the wall planes are detected and extracted using an efficient occlusion-aware wall surface segmentation technique [12], assuming that the room's interior is composed by vertical walls, which holds true for the vast majority of scenarios.

In the 2D projection of the wall plane points $\mathcal{P} = \{\boldsymbol{p}_0, \ldots \boldsymbol{p}_{N-1}\}$, their α-shape boundary $\mathcal{A}$ is computed, and refined in order to eliminate the introduced bias at borders and to improve the computed raw outline [18]. As a further refinement, we discard all open α-shape polylines and the closed ones that present a surface area less than $15cm^2$, assuming that the desired wall elements present ordinary architectural characteristics.

However, the α-shapes boundaries described above produce only a rough estimation of the true underlying object boundaries. In order to get regularized boundaries, a RANSAC algorithm is first used to robustly fit straight line segments to the refined α-shapes boundaries (see also Fig. 2(a)). Moreover, instead of keeping all estimated line models, the line space is further reduced to a set of mode (i.e. representative) line models using mean-shift clustering, and each line segment is then assigned to one of those modes (Fig. 2(b)).

Once all mode lines are detected, the occluded and actual empty areas in the wall surface have to be determined. Thus, similar to [1], we conduct an occupancy analysis on the wall surface in order to identify the clutter and non-clutter areas. In our approach, however, the sample points in a voxel are approximated by their centroid instead of the voxel center for accuracy reasons.

Next, a 2D cell complex representation $\mathcal{I}$ of the underlying wall surface is built. This cell complex is defined by all intersections of the representative mode lines in the wall plane and bounded by the wall's outline polygon as shown in Fig. 2.



(a) (b) (c)

Fig. 2: Cell complex construction. (a) Detected wall surface and the line models fitted to the (red) α-shapes outline. (b) Representative mode lines after mean-shift clustering. (c) 2D cell complex $\mathcal{I}$ (blue) imposed by the representative mode lines on the wall outline polygon.

3 Graph Cut Segmentation

Given the 2D cell complex $\mathcal{I}$, a weighted undirected dual graph G is defined by associating each cell c_p in $\mathcal{I}$ with a node in G and tagging each node by a semantic label from the set $\mathcal{L} = \{\mathcal{L}_{emp}, \mathcal{L}_{occ}\}$ indicating empty or occupied cells, respectively.

Using the Markov Random Fields (MRFs) and the graph-cut formulation [19], an optimal binary labeling can be found by minimizing the following objective function:

$$E(f) = \sum_{p \in V} \mathcal{D}_p(f_p) + \sum_{\{p,q\} \in N_G} V_{p,q}(f_p, f_q) \cdot T(f_p \neq f_q) \tag{1}$$

where $\mathcal{F} = \{f_0, \ldots f_{n-1}\}$ is a binary vector of assigned labels $f_i \in \mathcal{L}$ defining the segmentation, $\mathcal{D}_p(f_p)$ a unary data term, $V_{p,q}(f_p, f_q)$ a smoothness penalty function [20], N_G the 4-connected graph neighborhood, and $T(\cdot)$ is 1 if its argument is true and 0 otherwise.

In our framework, we additionally impose regional hard constraints for segmentation by identifying certain *seed* cells that should belong to wall surface or wall openings [19]. Thus we represent the wall cell complex $\mathcal{I}$ by a vector of n elements, $\boldsymbol{z} = [z_0(f_0), \ldots, z_{n-1}(f_{n-1})]$, indicating the estimated class probabilities $z_p(f_p) = Pr(f_p \mid c_p)$ of a cell c_p belonging to label f_p. This probability can be derived using the simplified Bayesian rule $Pr(f_p \mid c_p) \propto Pr(f_p)Pr(c_p \mid f_p)$, where the conditional probability $Pr(c_p \mid f_p)$ can be defined as

$$Pr(c_p \mid f_p) = \frac{\mid \varrho_p \mid}{K} e^{-|\mathcal{A}_p|/Cov_p}, \tag{2}$$

with K being a normalization factor, ϱ_p the density of points of cell c_p, $|\mathcal{A}_p|$ the number of α-extreme points in the cell's boundary line segments and Cov_p the number of points covering the cell c_p's interior.

For the computation of prior probability $Pr(f_p)$, we rely mainly on a distance function, which computes the normalized differences of cell densities. Specifically, let b_{occ} and b_{emp} be two labeled cells from $\mathcal{I}$, called *base* cells, for which we have a high certainty for their labels. Assuming that the cell with highest density ϱ belongs to the wall surface (label $\mathcal{L}_{occ}$), and the cell with lowest density belongs to a wall opening (label $\mathcal{L}_{emp}$), we assign these two (extreme) cells to be the two base cells. Further, we use the distance function $d(b_m, c_p)$, where $m \in \{occ, emp\}$, in order to estimate which cells from $\mathcal{I}$ are likely to have the same label as the associated base cell. Then, we compute the prior probability $Pr(f_p)$ as:

$$Pr(f_p) = n^{-1} \left|\{c_p \mid d(b_m, c_p) \leq 1\}\right|, \text{ when } f_p \equiv b_m \tag{3}$$

From $Pr(c_p \mid f_p)$ and $Pr(f_p)$ we estimate the class probabilities $Pr(f_p \mid c_p)$ and, based on two empirical constants κ_H and κ_L set to 0.9 and 0.2 respectively, we define the initial seed cell labelling as:

$$f_p = \begin{cases} \mathcal{L}_{occ} & \text{, if } z_p(\mathcal{L}_{occ}) \geq \kappa_H \\ \mathcal{L}_{emp} & \text{, if } z_p(\mathcal{L}_{emp}) \leq \kappa_L \end{cases} \tag{4}$$

Next, we compute the edge weights between the source s and sink t graph-cut terminals [19] and any remaining unlabelled graph nodes by

$$w_p^{(s)} = \beta \frac{z_p - \kappa_L}{\kappa_H - \kappa_L} \quad \text{and} \quad w_p^{(t)} = \beta \frac{\kappa_H - z_p}{\kappa_H - \kappa_L}, \tag{5}$$

where β is a sufficiently large constant ensuring a feasible flow in the graph, and we assign the t-links weights to the unary data term $\mathcal{D}_p(f_p)$ [19]:

$$\mathcal{D}_p(f_p) = w_p^{(s)} + w_p^{(t)} \tag{6}$$

Furthermore, we define edge weights used for the n-links as

$$w_{p,q} = \begin{cases} \kappa_n(\mathcal{R}_{dens} + \mathcal{R}_a) & \text{, if } p \neq q \\ 0 & \text{, if } p = q \end{cases}, \tag{7}$$

where $\mathcal{R}_{dens}$ and $\mathcal{R}_a$ are the contributions of certain cell features, and κ_n is a properly selected constant, which ensures that the n-link weights will be in the same range as the non-seed t-links.

The term $\mathcal{R}_{dens}$ is used to measure the difference of cell densities and is defined as

$$\mathcal{R}_{dens_{p,q}} = e^{-\frac{(\varrho_p - \varrho_q)^2}{2\sigma^2}}, \tag{8}$$

while the term $\mathcal{R}_a$ computes the difference of α-shape points between two neighboring cells

$$\mathcal{R}_{a_{p,q}} = \left||\mathcal{A}_p|^2 - |\mathcal{A}_q|^2\right|. \tag{9}$$

Given also these n-link weights, the binary labeling problem can be mapped to a graph cut problem, using Eq. 6 for the unary data terms and the smoothness penalty function [19] from Eq. 7:

$$V_{p,q}(f_p, f_q) = w_{p,q} \tag{10}$$

The solution of the global minimum cut with only two terminal nodes can then be computed from Eq. 1 in polynomial time [19].

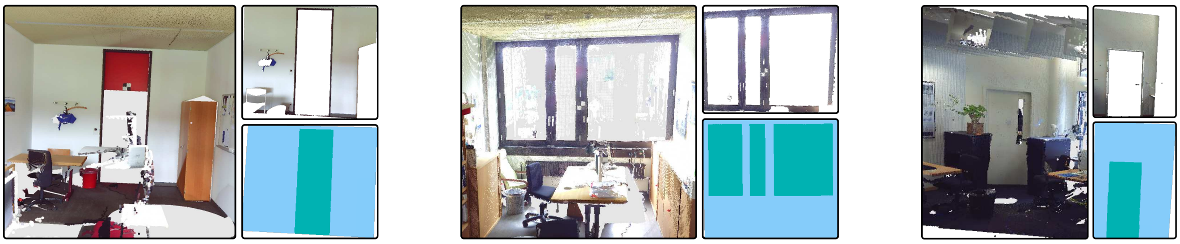
4 Experimental Results and Discussion

We have tested and evaluated our method on real-world datasets, which present different wall surfaces and occlusion levels. The selected rooms include windows and doors, and our algorithm is evaluated in both cases without manual tuning, as opposed to other techniques in the literature which are particularly tuned to identify only one of them (e.g. in [2,7,16]).

Fig. 3 shows the results of wall surface segmentation from different office environments. OFFICE1 dataset (Fig. 3(a)) contains an open doorway and a big occluder (wardrobe) in front of the wall. Our method was able to successfully reconstruct the door opening, without misclassifying the misleading opening from the occluder. OFFICE2 dataset (Fig. 3(b)) contains an irregular combination of window frames and several occluders, which hide a large part of the wall surface and the window frames too (in total, more than 45% of the reconstructed points were occluded). OFFICE3 dataset (Fig. 3(c)) contains a closed door along with some closets which cover part of it but also part from the door frame. Despite the challenging room environments, the wall openings were accurately detected in all three cases after the graph cut segmentation.

5 Conclusions

We have presented a method capable of reconstructing automatically the wall opening features in complex indoor environments. Our wall features segmentation pipeline was tested on real-world scenarios, proving that



(a) (b) (c)

Fig. 3: Wall surface reconstruction for (a) OFFICE1, (b) OFFICE2 and (c) OFFICE3. For each dataset, the input point cloud (left), the extracted wall plane (top right) and the wall openings segmentation (bottom right) are presented.

it can efficiently detect both windows and doors, as opposed to a majority of other current methods in the literature.

References

1. Antonio Adan and Daniel Huber. 3D reconstruction of interior wall surfaces under occlusion and clutter. *3D Imaging, Modeling, Processing, Visualization and Transmission*, pages 275–281, May 2011.
2. L. Díaz-Vilariño, J. Martínez-Sánchez, S. Lagüela, J. Armesto, and K. Khoshelham. Door recognition in cluttered building interiors using imagery and lidar data. *International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences*, XL-5:203–209, June 2014.
3. Xiong Xuehan, Adan Antonio, Akinci Burcu, and Huber Daniel. Automatic creation of semantically rich 3d building models from laser scanner data. *Automation in Construction*, 31:325–337, January 2013.
4. Arayici Yusuf. An approach for real world data modelling with the 3D terrestrial laser scanner for built environment. *Automation in Construction*, 16(6):816–829, 2007.
5. Yasutaka Furukawa, Brian Curless, Steven M. Seitz, and Richard Szeliski. Manhattan-world stereo. In *Proceedings IEEE Conference on Computer Vision and Pattern Recognition*, pages 1422–1429, June 2009.
6. Brian Okorn, Xuehan Xiong, Burcu Akinci, and Daniel Huber. Toward automated modeling of floor plans. In *Proceedings Symposium on 3D Data Processing, Visualization, and Transmission*, volume 2, pages 17–20, May 2010.
7. Angela Budroni and Jan Böhm. Automatic 3D modelling of indoor manhattan-world scenes from laser data. *International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences*, XXXVIII:115–120, February 2010.
8. Daniel Huber, Burcu Akinci, Antonio Adan Oliver, Engin Anil, Brian E. Okorn, and Xuehan Xiong. Methods for automatically modeling and representing as-built building information models. In *Proceedings NSF CMMI Research Innovation Conference*, January 2011.
9. Sven Oesau, Florent Lafarge, and Pierre Alliez. Indoor scene reconstruction using feature sensitive primitive extraction and graph-cut. *ISPRS Journal of Photogrammetry and Remote Sensing*, 90:68–82, 2014.
10. Pingbo Tang, Daniel Huber, Burcu Akinci, Robert Lipman, and Alan Lytle. Automatic reconstruction of as-built building information models from laser-scanned point clouds: A review of related techniques. *Automation in Construction*, 19(7):829–843, June 2010.
11. Rebekka Volk, Julian Stengel, and Frank Schultmann. Building information modeling (BIM) for existing buildings - Literature review and future needs. *Automation in Construction*, 38:109–127, March 2014.
12. Claudio Mura, Oliver Mattausch, Alberto Jaspe Villanueva, Enrico Gobbetti, and Renato Pajarola. Automatic room detection and reconstruction in cluttered indoor environments with complex room layouts. *Computers and Graphics*, 44:20–32, 2014.
13. Jianxiong Xiao and Yasutaka Furukawa. Reconstructing the world's museums. *International Journal of Computer Vision*, 110(3):243–258, 2014.
14. Alexander Hinneburg and Daniel A. Keim. An efficient approach to clustering in large multimedia databases with noise. In *International Conference on Knowledge Discovery and Data Mining*, pages 58–65, 1998.
15. Remus-Claudiu Dumitru, Dorit Borrmann, and Andreas Nüchter. Interior reconstruction using the 3D Hough Transform. *International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences*, XL-5/W1:65–72, February 2013.
16. Richard Zhang and Avideh Zakhor. Automatic identification of window regions on indoor point clouds using LiDAR and cameras. *IEEE Applications of Computer Vision*, pages 107–114, March 2014.
17. Mattia Previtali, Marco Scaioni, Luigi Barazzetti, and Raffaella Brumana. A flexible methodology for outdoor/indoor building reconstruction from occluded point clouds. *Annals of Photogrammetry, Remote Sensing and Spatial Information Sciences*, 2(3):119–126, September 2014.
18. Shen Wei, Yunhao Chen, Li Jing, Siti Atikah, and Jalan Tun Ismail. Rebuilding the 3D models of buildings based on lidar data. In *Asian Conference on Remote Sensing*, volume I, pages 995–1003, November 2007.
19. Yuri Boykov and Gareth Funka-Lea. Graph cuts and efficient N-D image segmentation. *International Journal of Computer Vision*, 70(2):109–131, November 2006.
20. Sara Vicente, Vladimir Kolmogorov, and Carsten Rother. Graph cut based image segmentation with connectivity priors. *IEEE Computer Vision and Pattern Recognition*, pages 1–8, June 2008.